

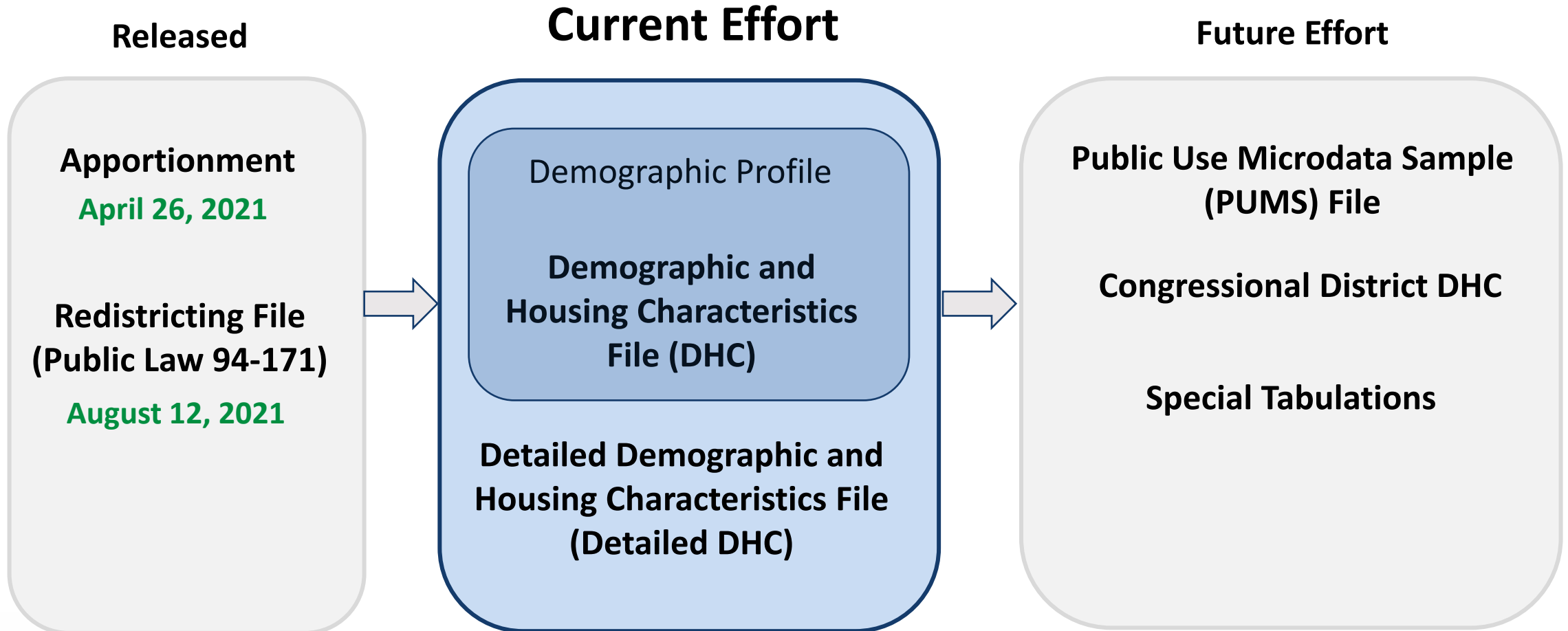
# Update on the Demographic Profile and Demographic and Housing Characteristics File (DHC)

Jason Devine  
Population Division

Matthew Spence  
Population Division

**National Advisory Committee 2022 Spring Virtual Meeting**  
**May 5-6, 2022**

# 2020 Census Data Products



# Topics Included in the 2020 Census

- Relationship to Householder
- Sex
- Age
- Hispanic or Latino origin
- Race
- Housing Tenure (owned, rented, occupied without payment)
- Vacancy Status

# The DHC and Demographic Profile

## Demographic and Housing Characteristics File (DHC)

Includes many of the demographic and housing tables previously included in the 2010 Census Summary File 1. In addition to demographic and housing tables the DHC will also include counts of people by relationship to the householder. The DHC will include counts of households or families by type, but it will not include counts of people by those household or family types. Geography varies with many tables provided at the Census Block level.

## Demographic Profile

Provides critical demographic and housing characteristics about local communities from selected DHC tables in a streamlined, easy to use format. In previous decades, this product was released as soon after the release of the Redistricting product as possible. With the implementation of the new disclosure avoidance methodology the production of the Demographic Profile tables follows the same timeline as the DHC. The current plans include for the release of the Demographic Profile at the same time as the DHC.

# Determining the Contents of the DHC

- The initial proposal for the design and contents of the DHC was developed and shared in 2019 based on subject matter expert knowledge and feedback from a Federal Register Notice
- The 2010 Census Summary File 1 (SF1) was the starting point for the design of the DHC with the DHC including many of the tables included in SF1
- The contents of the DHC were reviewed and updates proposed based on the feedback received documenting uses of the data
- The final contents of the DHC will be reflected in the second round DHC demonstration product

# Some Topics Included in the DHC

## **Counts of People by:**

- Sex by single year-of-age
- Hispanic or Latino origin by race by sex by age
- GQ population by sex by age
- Household type by relationship to householder

## **Counts of Households/Families by:**

- Multigenerational households
- Hispanic or Latino origin of householder by race of householder
- Family type by presence of children
- Presence of people of different ages
- Housing tenure (owner, renter occupied) by characteristics of the householder or household
- Vacancy status

# DHC Demonstration Products

- Demonstration products allow the public to see progress and the level of noise that is being added through comparisons between the published 2010 Census data and 2010 Census data protected using the new disclosure avoidance methodology
  - 5 demonstration products and a production settings product were provided leading up to the release of the redistricting data (2 of these included limited DHC content, the other 3 only included redistricting content)
  - Two DHC specific demonstration products (including the 2022-03-16 demonstration product) and a production settings product are planned for the DHC

Previous  
Demonstration  
Products

## ^ About DHC Development

The census data product formerly known as "Summary File 1" is now the Demographic and Housing Characteristics File (DHC). Please follow our [newsletter](#) for updates and to learn how to contribute to its development. Review the [2020 Census Data Product Planning Crosswalk](#) to better understand the proposed changes to this data product.

### ∨ DAS Development Update 2022-03-16

[Top of Section](#)

## Detailed Demographic and Housing Characteristics File (Detailed DHC)

### ^ About Detailed DHC Development

Improvements in the design, processing and coding of the 2020 Census allow the release of data for almost five times as many detailed race and ethnic groups than were possible in 2010. Development of disclosure protections to accommodate this improvement are underway. See the "Detailed DHC" tabs in the [2020 Census Data Product Planning Crosswalk](#).

[Top of Section](#)

## 2020 Census Redistricting Data (P.L. 94-171)

[EXPAND ALL](#) | [COLLAPSE ALL](#)

### ∨ About 2020 Census Redistricting Data (P.L. 94-171) Summary File Development

### ∨ Production Settings: 2021-06-08

### ∨ DAS Development Update 2021-04-28

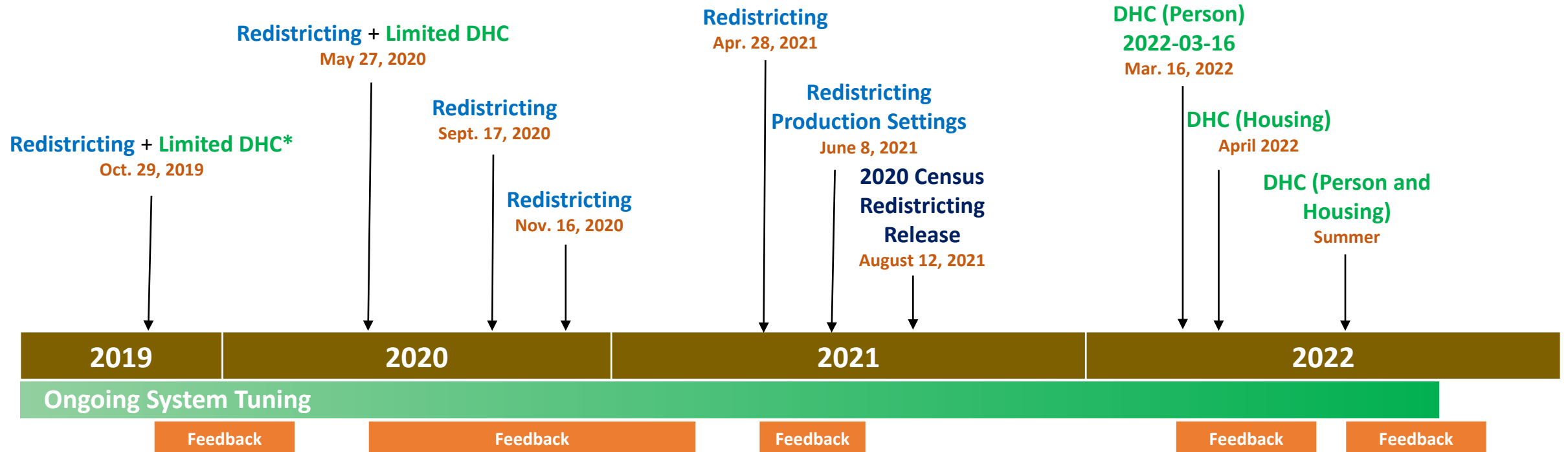
### ∨ DAS Development Update 2020-11-16

### ∨ DAS Development Update 2020-09-17

### ∨ DAS Development Update 2020-05-27

### ∨ 2010 Demonstration Data Products Baseline 2019-10-29

# Demonstration Product Timeline





# Refining Our Disclosure Avoidance Methodology

- Experimental runs from our Disclosure Avoidance System (DAS) using 2010 Census data allow us to assess the resulting privacy and accuracy
- A re-identification study is being used to assess the risk of re-identification for a specific attack
- Leading experimental run along with the results for the accuracy analysis and re-identification study are discussed with DSEP\* to obtain approval to release as the demonstration product
- The demonstration product is made available to the public for their analysis with their comments helping guide work on improvements to the DAS

\* Data Stewardship Executive Policy Committee

# The DHC Demonstration Products

- Two rounds of DHC demonstration products are planned
- The 2010 Demonstration Data Product for the DHC (2022-03-16) for tables sourced from the **Person** File in the form of Summary Files was released for public analysis on March 16
- Accuracy metrics for the tables sourced from both the **Person** and **Housing Unit (Unit)** file demonstration products were also released on March 16
- A DHC demonstration product for tables sourced from the **Unit** file was released on March 29
- The public has 30 days from the release of the **Unit** file to review and comment on both the **Person** and **Unit** files
- An issue in the Unit Summary Files was identified and corrected, the review period was extended to May 16 to account for this
- This is the most extensive set of demonstration data to be released using the new disclosure avoidance methods (113 Person File tables, 121 Unit File tables, 50+ different geographies)

# Receiving and Using Feedback on the DHC

- Comments received on the Redistricting data demonstration products informed the tuning of the DAS and the setting of the privacy loss budget for that product
- **Over 400** comments on the contents of the DHC and Detailed DHC were received throughout the October and December 2021 comment periods
- Most uses that require basic demographic and housing information will be broadly met by the tables currently proposed for the DHC
- Analysis is being conducted to determine if a set of tables can be added back to the DHC to better meet the needs of the uses documented in the comments
- While many comments expressed a need for Census block-level data some tables will still not be available at that level because of the impact on privacy and accuracy

# Receiving and Using Feedback on the DHC (cont.)

- For the first round of DHC demonstration data we wanted feedback on the accuracy of the tables and the impact this has on uses of the data
- We have continued to work to make improvements based on our own internal analysis
- We are in the process of reviewing the comments we received on the first round of DHC demonstration data so they can be considered in the work leading up to the second round of DHC demonstration data
- Considering the comments on the first round of DHC demonstration data and ongoing improvements informed by internal analysis we will use comments on the second round of DHC demonstration data to inform final improvements before DSEP\* establishes the DHC Disclosure Avoidance System (DAS) production settings

\* Data Stewardship Executive Policy Committee

# Evaluating DHC Demonstration Data

## Detailed Summary Metrics

- Create measures of **accuracy**, **bias**, and **outliers** based on tabulated quantities at various geographic levels (e.g., Total Population at the county level, Asian Alone at the tract level) for a DAS run to published (i.e., swapped) 2010 tabulations

## Tabulations / Maps

- Directly compare values using tables or maps to illustrative substantive accuracy/error

# Detailed Summary Metrics

Accuracy	<ul style="list-style-type: none"><li>- Mean Absolute Error (MAE): What is the average absolute change (+/-)?</li><li>- Mean Absolute Percent Error (MAPE): What is the average relative change (+/- %)</li></ul>
Bias	<ul style="list-style-type: none"><li>- Mean Algebraic Percent Error (MALPE): What is the average directional relative change?</li></ul>
Outliers	<ul style="list-style-type: none"><li>- How many geographies are above particular thresholds?</li></ul>

## Pros:

- Quick way to summarize across multiple experimental runs or over time
- Does not require additional data wrangling or tabulation

## Cons:

- Relative error can be misleading for small population sizes
- Does not communicate *substantive* importance

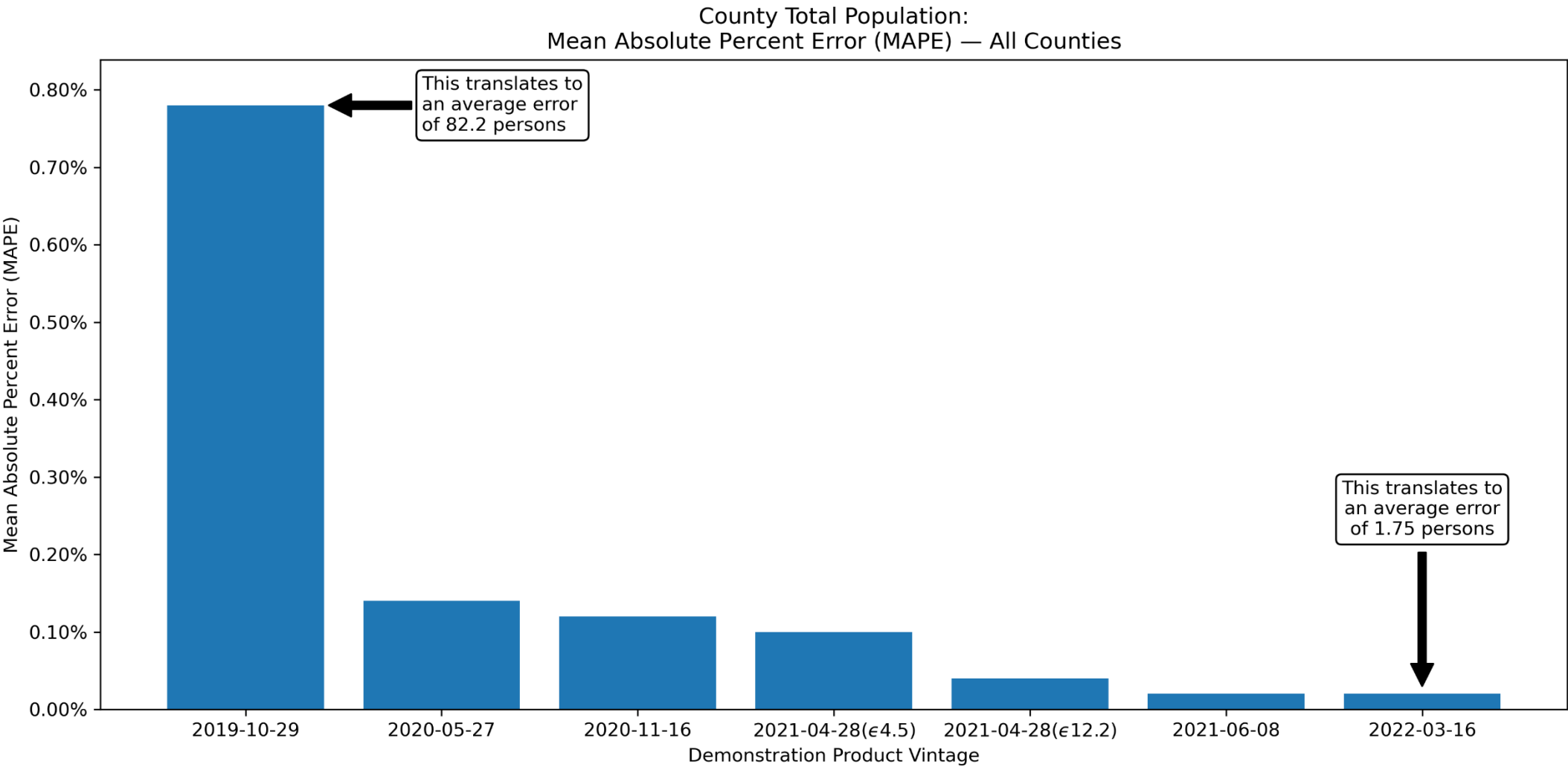
# Calculating Metrics: MAE

County	Published 2010 Population	PPMF 2010 Population	Error	Absolute Error
Autauga County, Alabama	54,571	54,574	3	3
Baldwin County, Alabama	182,265	182,266	1	1
Barbour County, Alabama	27,457	27,456	-1	1
Bibb County, Alabama	22,915	22,917	2	2
Blount County, Alabama	57,322	57,322	0	0
...	...	...	...	...
Loving County, Texas	82	85	3	3
...	...	...	...	...

Mean Absolute Error: 1.75



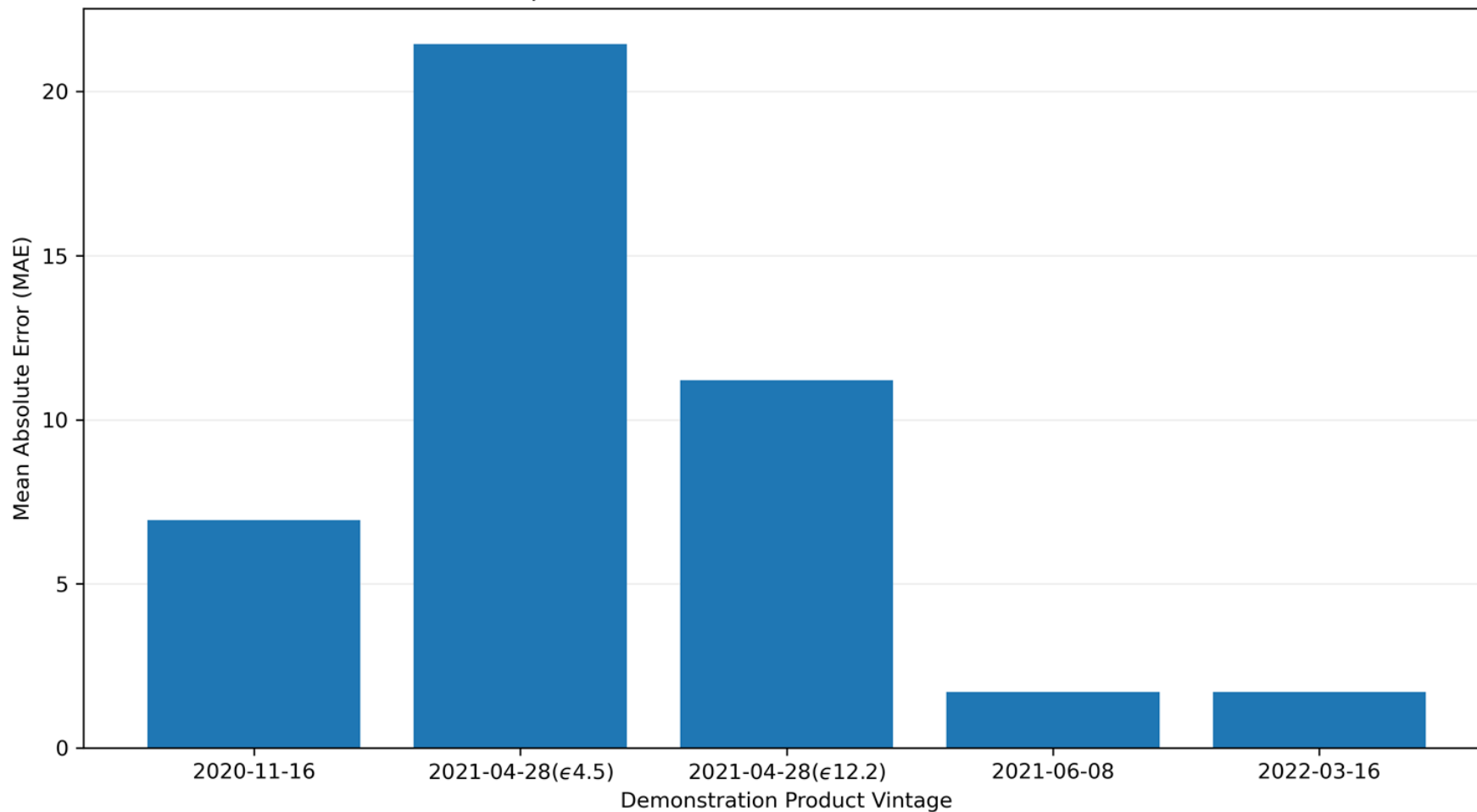
# Total Population – County



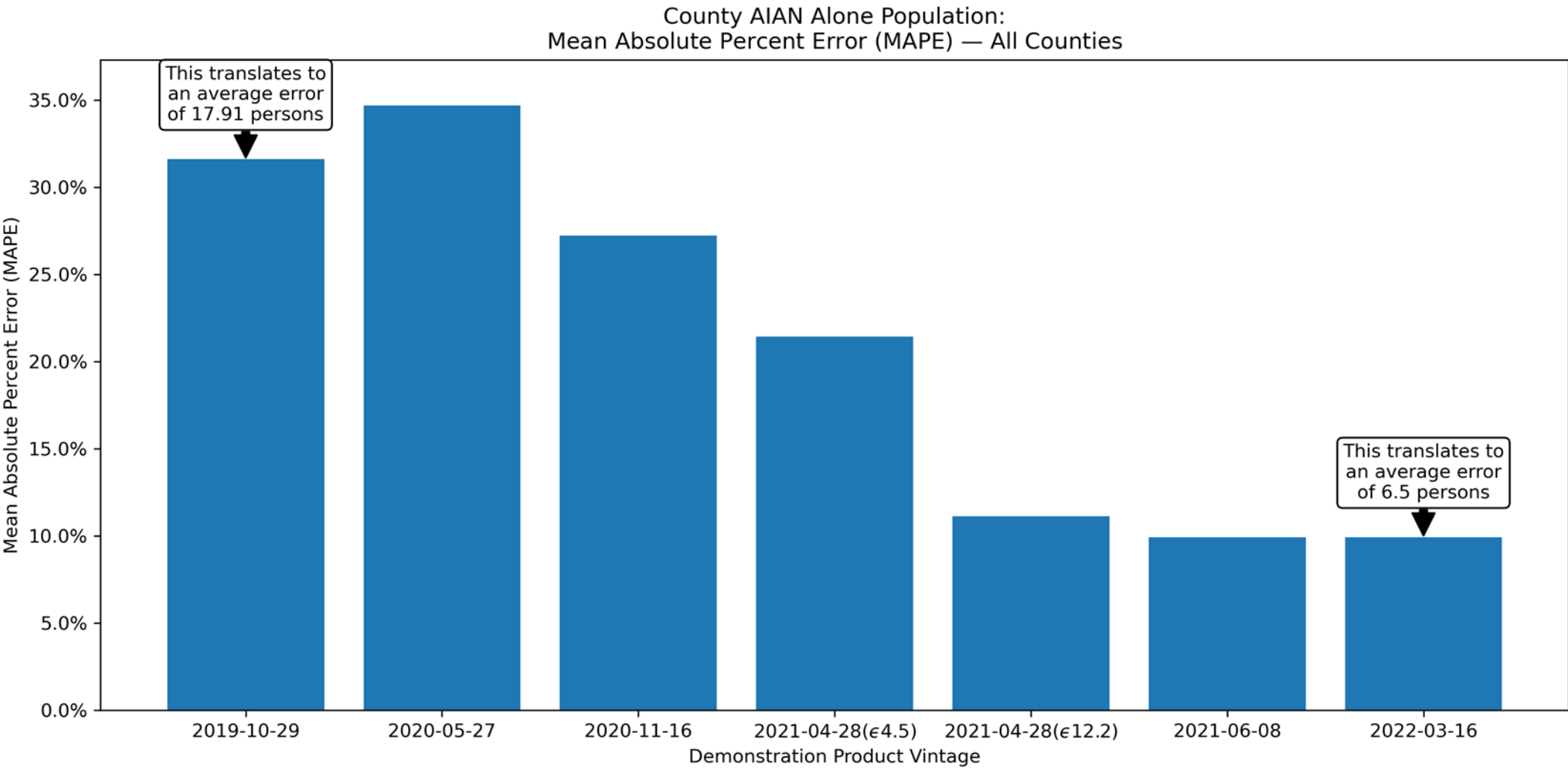


# Total Population – Federal American Indian Reservation

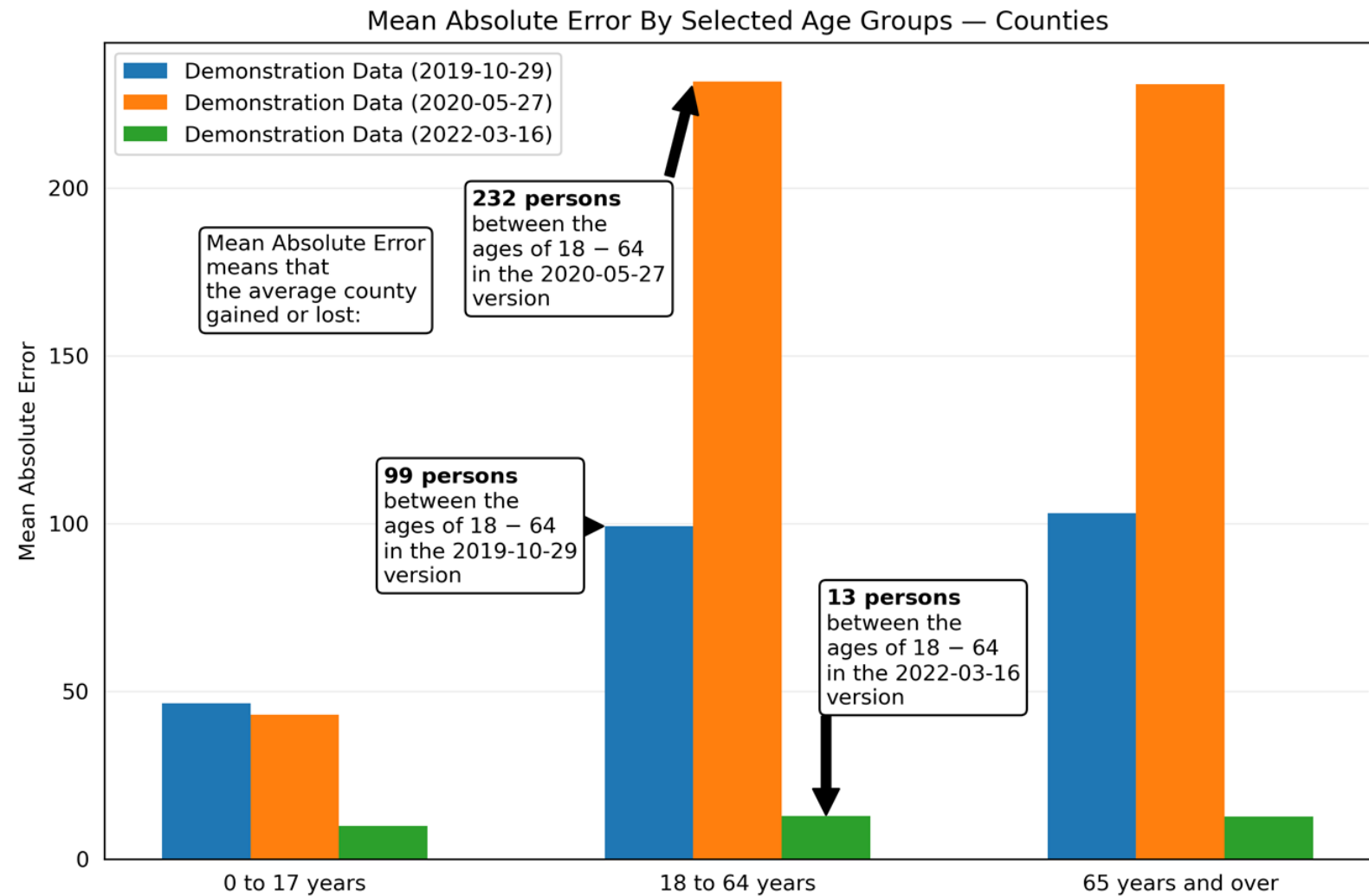
Federal American Indian Reservation/Off-Reservation Trust Land  
Total Population: Mean Absolute Error (MAE) — All Areas



# Population of AIAN Alone – County



# Population by Sex and Selected Age – County



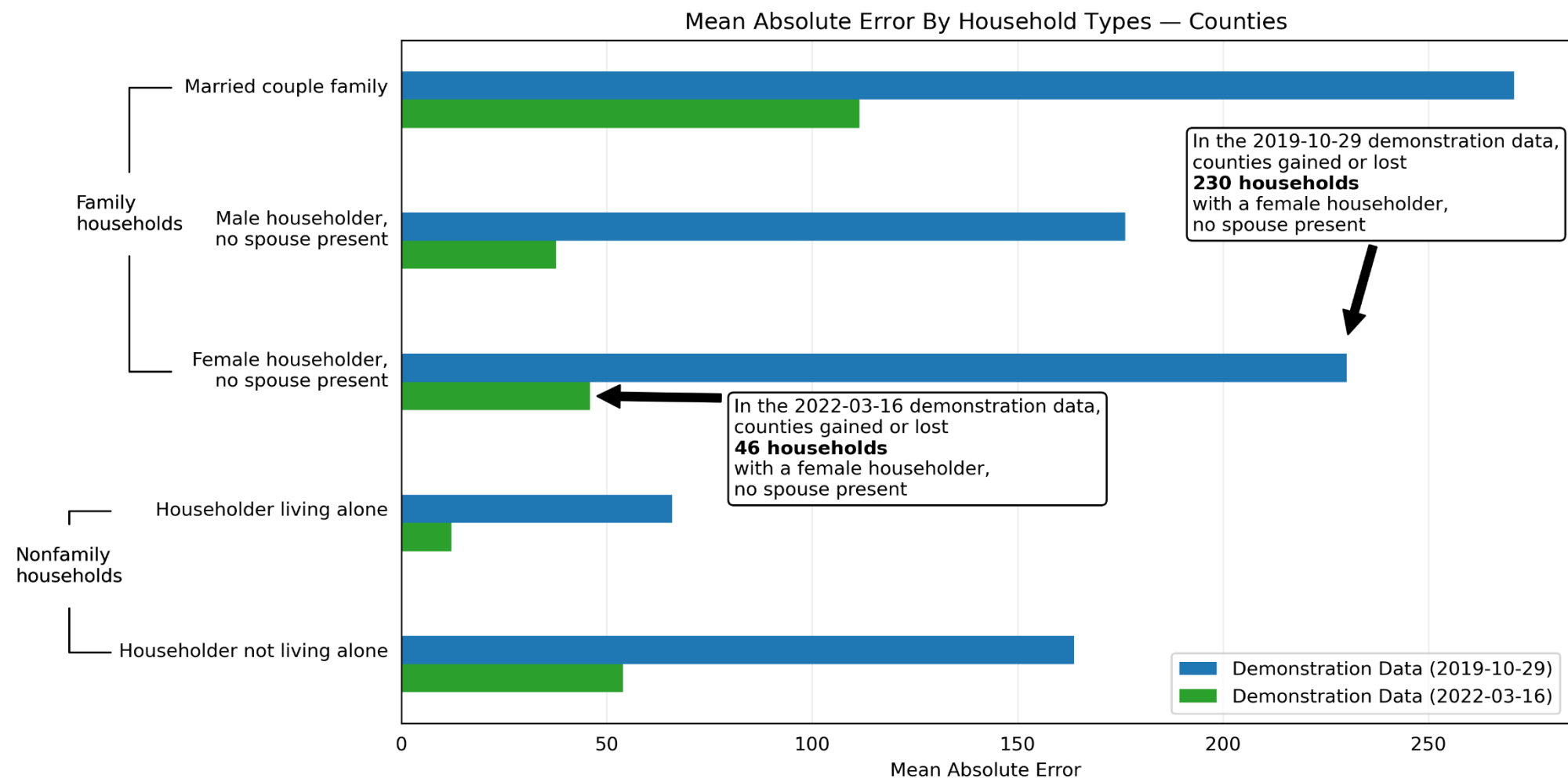
# Mean Absolute Error for 5 Year Age Groups – Counties

	Demonstration Data (2019-10-29)	Demonstration Data (2020-05-27)	Demonstration Data (2022-03-16)
Under 5 years	96.50	83.67	<b>7.75</b>
5 to 9 years	111.47	81.73	<b>7.42</b>
10 to 14 years	110.19	81.50	<b>7.36</b>
15 to 19 years	100.38	92.31	<b>12.03</b>
20 to 24 years	100.12	106.48	<b>18.20</b>
25 to 29 years	107.86	95.39	<b>16.06</b>
30 to 34 years	114.02	89.38	<b>12.53</b>
35 to 39 years	97.79	87.08	<b>12.17</b>
40 to 44 years	97.65	84.14	<b>12.07</b>
45 to 49 years	116.99	87.00	<b>11.95</b>
50 to 54 years	110.95	84.56	<b>11.17</b>
55 to 59 years	97.19	82.07	<b>10.93</b>

# Mean Absolute Error for 5 Year Age Groups – Counties

	Demonstration Data (2019-10-29)	Demonstration Data (2020-05-27)	Demonstration Data (2022-03-16)
60 to 64 years	99.92	81.74	<b>6.76</b>
65 to 69 years	99.45	76.22	<b>7.02</b>
70 to 74 years	99.40	76.12	<b>9.69</b>
75 to 79 years	85.14	72.49	<b>9.43</b>
80 to 84 years	80.09	69.43	<b>8.87</b>
85 to 89 years	78.29	65.74	<b>23.63</b>
90 to 94 years	65.03	53.44	<b>16.83</b>
95 to 99 years	39.15	35.26	<b>14.71</b>
100 to 104 years	10.77	12.86	<b>6.46</b>
105 to 109 years	3.71	2.80	<b>2.00</b>
110 to 115 years	4.27	2.43	<b>2.24</b>

# Population by Sex and Selected Age – County



Demonstration Data (2020-05-27) did not include tables sourced from the unit file

# Tabulations & Maps

Directly tabulate variables of interest and compare them to published 2010 tables (SF1).

Pros:

- Can create custom tabulations for a particular area or value of interest
- Clear sense of substantive importance of error – how much change for this geography?
- Can create maps to see spatial patterns

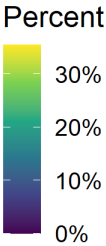
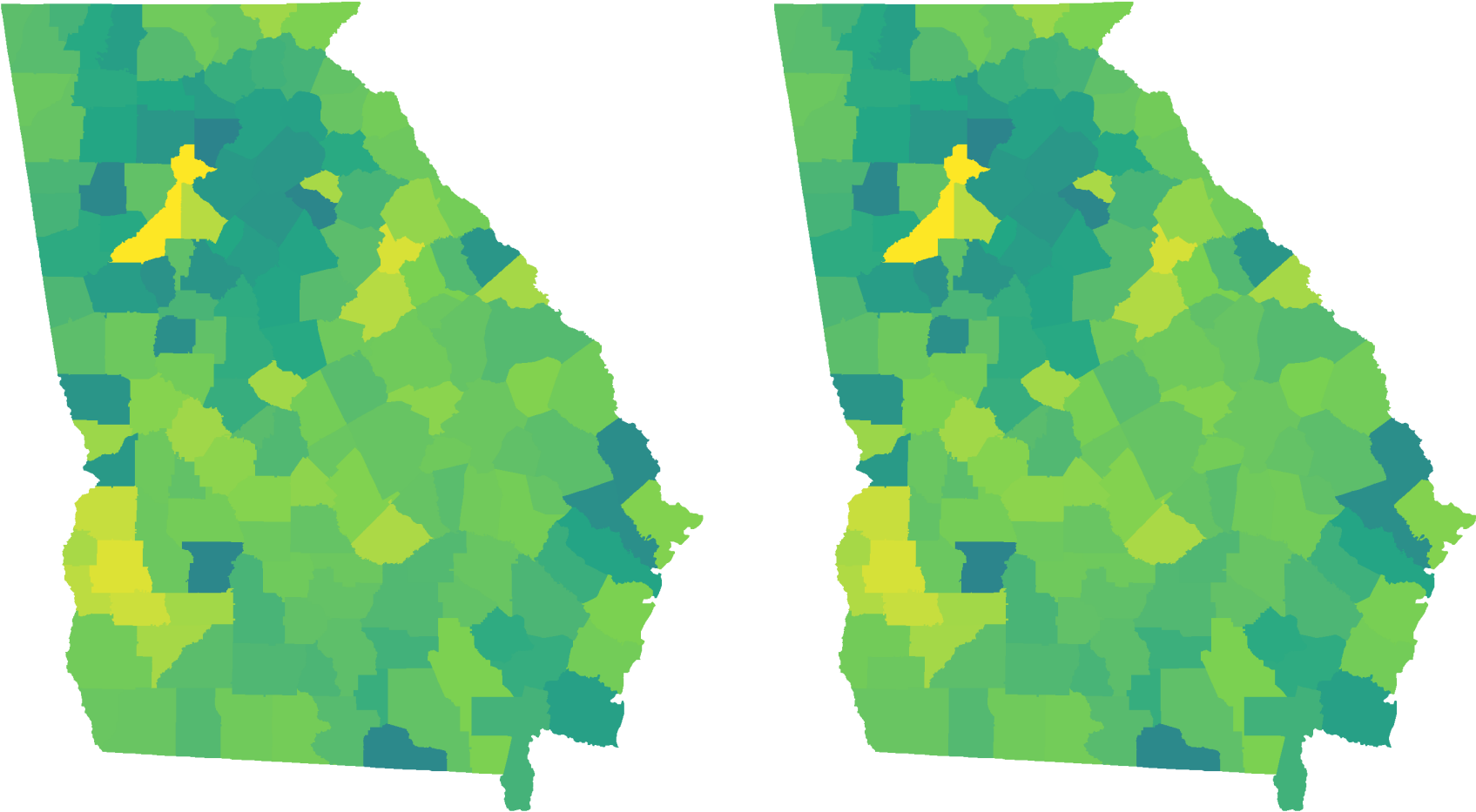
Cons:

- Must use statistical software to read in DHC Demonstration Data.
- Must be familiar with technical documentation (e.g., geographic component)
- Tables may have changed sourcing or universe (crosswalk very important here)
- Certain variables (e.g., foster children) did not exist on the 2010 published tables

# Share of Householders Living Alone – County

SF1 (Published)

2022-03-16 Demonstration Product

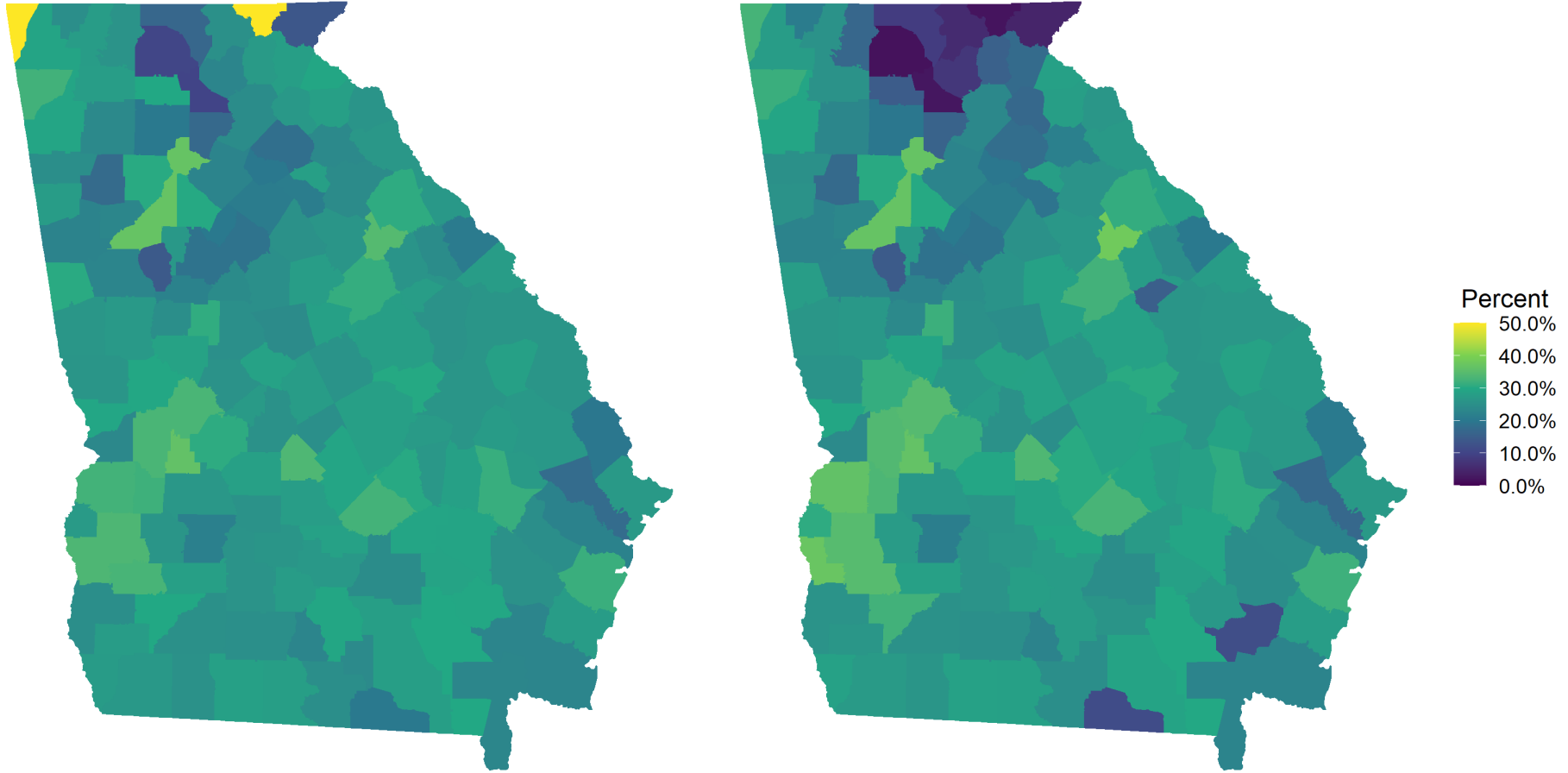




# Share of Householders Living Alone – Black or African American Alone Householder – County

SF1 (Published)

2022-03-16 Demonstration Product



# Questions?



Please send your comments on the demonstration data to [2020DAS@census.gov](mailto:2020DAS@census.gov)

# Questions for NAC

- What questions or concerns do you have about the DHC demonstration products?
- What questions or concerns do you have about the content planned for the DHC?
- Do you have any questions or concerns about any of the information presented today?